Chapter 2 Fundamentals of Ocular Wavefront Correction

For vision correction, the key is to help patients achieve good visual performance. In most cases, this means good visual acuity and good contrast sensitivity. The ocular aberrations can be considered as a combination of the spherocylindrical aberrations and irregular aberrations. In wavefront optics terminology, they are called low-order and high-order aberrations, respectively. In general, the degradation of visual performance due to the low-order aberrations is much higher than that due to the high-order aberrations. Traditional vision correction thus focuses on the correction of low-order aberrations.

However, to achieve a better standard of vision, it is not enough just to correct low-order aberrations. For example, professional shooting athletes and fighter pilots may seek supervision to enhance their performance. The correction of higher order aberrations, however, is not limited to these small groups. The average person may seek enhanced vision to provide a reserve for ageing vision or to eliminate night-vision problems, such as halos, glares, etc.

2.1 Principle of Phase Conjugation

Degradation of celestial images due to astronomical seeing has been a huge problem for centuries in astronomical observations. In 1953, Bobcock proposed a solution that has evolved into an important technology called adaptive optics.[1] The major concept in adaptive optics is phase conjugation,[1] where optical aberrations can be compensated by a deformable mirror. Today, applications of adaptive optics can be found in astronomy,[2] the military,[3] and vision.[4]

This section discusses some of the basic concepts, such as the wavefront and the optical path difference, as well as their use in phase conjugation for the purpose of vision correction.

[1] In general, phase conjugation is a phenomena in nonlinear optics representing a phase reversal of a light beam after reflecting from a phase-conjugating mirror. This terminology was borrowed by the adaptive optics community to represent a phase compensation or cancellation between a propagating wavefront and a deformable mirror.

2.1.1 Refractive Index of the Cornea

In vision, light passes into the human eye and reaches the retina through which information is perceived by the brain. The human eye as an optical system consists of a few layers of transparent media. Among them, the cornea and the crystalline lens are the two most important components. Different kinds of media have different effects on the propagation of the light. For the purpose of corneal ablation, the index of refraction n is the most important property.

The cornea consists of five layers: epithelium, Bowman's membrane, stroma, Descemet's membrane, and endothelium. The epithelium covers the surface of the cornea with five to six layers of cells. It is about 50 microns thick and can regenerate in a few days if injured. Bowman's membrane is a very thin (about 5 to 10 μm) layer of condensed, irregularly-arranged collagen fibers to protect the stroma from damage. The stroma is a thick layer of about 550 microns consisting of approximately 200 layers of collagen fibrils. The stroma does not regenerate if damaged. Laser ablation aims to reshape the corneal stroma for vision correction. Descemet's membrane is a thin acellular layer serving as the basement for the corneal endothelium. The endothelium is a monolayer of cells at the posterior cornea.

The indices of refraction of the epithelium and the anterior and posterior surfaces of the stroma at visible light[2] were measured to be 1.401, 1.380, and 1.373, respectively.[5] For refractive surgery, where laser ablation is applied to the stroma, the refractive index of the stroma is of major importance. In this book, we use the average of the indices of refraction between the anterior and posterior surface of the stroma, 1.3765, as the mean value of the refractive index of the cornea.

2.1.2 Wavefront and Optical Path Difference

In physical optics, a wavefront is related to the phase of the complex field of an eletromagnetic wave. In geometrical optics, a wavefront is an envelope of rays with the same phase. Analogous to water waves, the wavefront peak represents a phase advance, and the wavefront valley represents a phase lag. When an incoherent light passes through two different routes as two rays, they may not reach the destination at the same time. This time difference between the two rays, translated into the distance by multiplying the time difference by the speed of the light in a vacuum, results in the optical path difference.

Figure 2.1 shows an example of two rays from a plane wave: AB and OP. Ray AB passes through the edge of the lens, and ray OP passes through the center of the lens. Ray AB takes $t_1 = l/c$, where c is the speed of the light. Because the light travels slower in the glass, ray OP takes $t_2 = nl/c$. Hence, the optical path difference between ray AB and ray OP is $(t_2 - t_1)c = (n-1)l$. Consequently, ray OP is slower than ray AB, resulting in a phase lag.

In this book, we assign a positive value for the optical path difference to a phase

[2]The index of refraction of any medium depends upon the wavelength of the incident light. For visible light, or white light, 550 nm is often used as the wavelength λ.

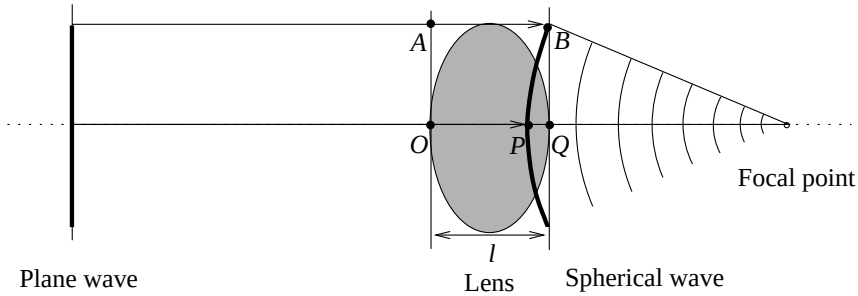


Figure 2.1: Calculation of the optical path difference for a thin lens with refractive index of n. The physical thickness of the lens is l, so the ray that travels through the center of the lens has an optical path length of nl. On the other hand, the ray that travels through the edge of the lens has an optical path length of exactly l. Hence, the optical path difference is $(n-1)l$.

advance and a negative value to a phase lag. In the previous example, therefore, the phase value at point B is 0, and that at point Q is negative. So, rays going through the periphery of the lens bend more, eventually hitting the point of focus. Similarly, for a diverging wavefront, the phase value at the optical axis is more positive.

Example 2.1: If the medium in Fig. 2.1 is corneal stroma, and $l = 20$ μm, calculate the OPD, or the phase lag, between point B and point Q.
Solution: The optical path difference is $(n-1)l = (1.3765-1) \times 20 = 7.53$ μm. Note that the OPD for stroma is 0.3765 times its physical depth. In other words, the cornea depth is about 2.656, or exactly $1/(n-1)$, times the OPD.

If the medium is not homogenous, as in the case in Fig. 2.2, the optical path lengths of different points relative to a reference surface are different. We can connect all points in space that have the same optical path length to form a wavefront. In other words, a wavefront is the surface of points having the same phase.

2.1.3 Phase Conjugation

Phase conjugation is the key concept in adaptive optics. It is also the basis for wavefront-driven refractive surgery. It indicates that the optical path difference of the correcting optics matches the optical path difference of the wavefront to remove the optical aberrations. Figure 2.3 explains how phase conjugation works. For a plane wave propagating towards a flat mirror, the reflected wavefront is the same as the impinging wavefront, as the mirror does not change the optical path length but only reverses the propagation direction. Similarly, a normal wavefront becomes inverted due to the reflection from the mirror but otherwise remains the same shape. However, when the flat mirror is changed to a deformable mirror with half the magnitude of the impinging wavefront, the reflected wavefront becomes flat. When a normal wavefront

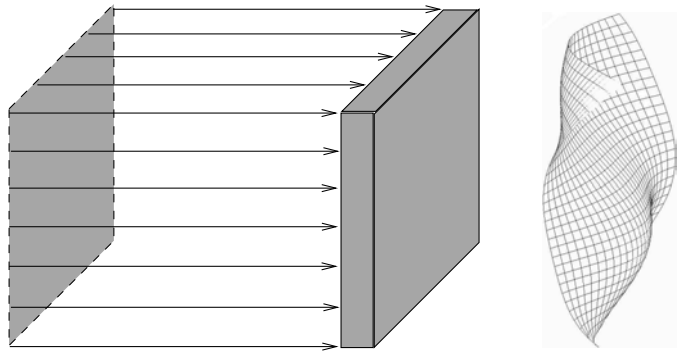


Figure 2.2: A plane wave, propagating through a nonhomogenous medium with a varying index of refraction, has the same phases at different points. The surface connecting all these points forms a wavefront.

passes through a deformable lens with the same optical path length of the impinging wavefront but with a different sign, the refracted wavefront also becomes flat. The physical shape of the front surface of the lens is $1/(n-1)$ times the shape of the impinging wavefront in order to cancel the optical path difference of the impinging wavefront.

In refractive surgery, we apply the phase conjugation as in Fig. 2.3 (d). The idea is to change the shape of the cornea so that the optical path difference of the total ocular aberrations among different points of the cornea can be compensated by the topographic change of the cornea. For most refractive surgery techniques, where the laser ablation applies to the stroma, the treatment profile of the stroma tissue must be scaled by a factor of $1/(n-1) = 2.656$ from the ocular aberrations, measured by a wavefront aberrometer.

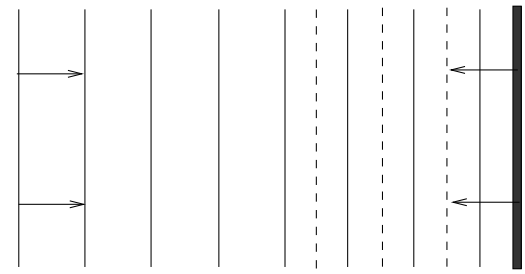
2.2 Munnerlyn Equation

Traditionally, nearsightedness (myopia), farsightedness (hyperopia), and astigmatism can be corrected with spectacles and contact lenses. They can also be corrected with laser refractive surgery. In this section, we will discuss vision correction with laser refractive surgery, from the conventional spherocylindrical correction to customized wavefront correction.

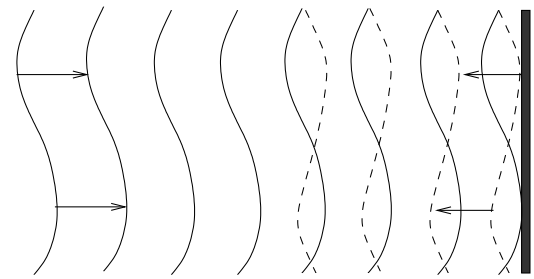
Finding a treatment profile for refractive surgery was first investigated by Munnerlyn and his associates.[6] Therefore, the formula that described the ablation profile is often called the Munnerlyn equation.

2.2.1 Myopic Correction

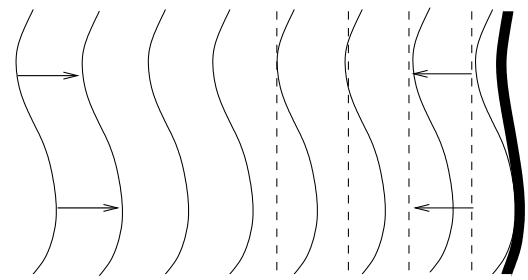
Figure 2.4 shows the schematic for the design of myopic and hyperopic ablation profiles using thin lens optics. It can be shown (Appendix 2.A) that the Munnerlyn equation for myopic correction is given by



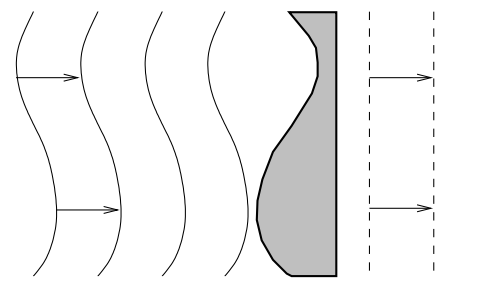
(a)



(b)



(c)



(d)

Figure 2.3: Principle of the phase conjugation. (a) a plane wave propagating toward a flat mirror; (b) a normal wavefront propagating toward a flat mirror; (c) a normal wavefront propagating toward a deformable mirror; (d) a normal wavefront propagating through a deformable lens with a sign reversal.

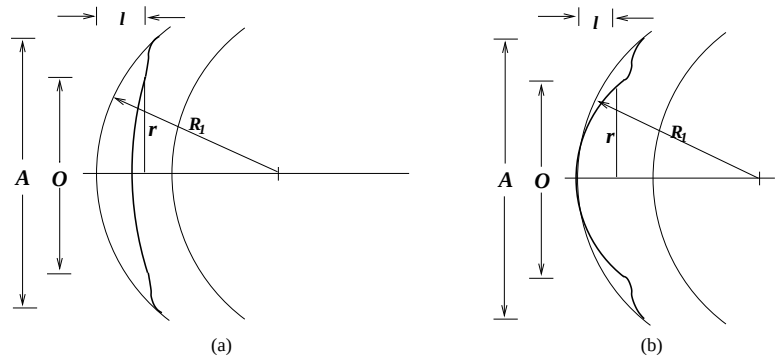


Figure 2.4: Geometry of refractive correction profiles for the derivation of the Munnerlyn equation. (a) Myopic; (b) Hyperopic.

$$\begin{aligned} l(r) &= \sqrt{R_1^2 - r^2} - \sqrt{\left[\frac{(n-1)R_1}{n-1+SR_1}\right]^2 - r^2} \\ &\quad + \sqrt{\left[\frac{(n-1)R_1}{n-1+SR_1}\right]^2 - (O/2)^2} - \sqrt{R_1^2 - (O/2)^2}, \end{aligned} \tag{2.1}$$

where r is the distance of a point to the center of the pupil in meters, R_1 is the radius of curvature of the front surface of the cornea in meters, n is the refractive index of the ablation material (i.e., the stroma), S (the denoting sphere) is the myopic (hence negative) refraction in diopters, and O is the diameter of the optical zone in meters.

The ablation profile, as determined by the Munnerlyn equation, is exact as defined by Eq. (2.1). However, the actual ablation zone that the ablation is applied to is larger than the optical zone. For a myopic eye with zero astigmatism, the ablation zone can be as large as the optical zone. But for myopic astigmatic eyes, different meridians have different ablation depths, so a transition zone is needed to bring the different heights along different meridians within the optical zone to the zero level at the periphery of the ablation zone (indicated by A in Fig. 2.4).

Example 2.2: Calculate the tissue ablation depth for a -1 D myopic eye using a 6 mm optical zone.

Solution: From Eq. (2.1), set $r = 0$, $S = -1$ D, $O = 0.006$, and use $R_1 = 7.8$ mm, we obtain $l(0) = 0.0078 - \frac{(1.3765-1)\times 0.0078}{1.3765-1-1\times 0.0078} + \sqrt{\left[\frac{(1.3765-1)\times 0.0078}{1.3765-1-1\times 0.0078}\right]^2 - (0.006/2)^2} - \sqrt{0.0078^2 - (0.006/2)^2} = 13.43 \times 10^{-6}$ m $= 13.43$ μm. Hence, for every diopter myopia, the ablation depth is about 13.5 μm.

2.2.2 Hyperopic Correction

For hyperopic ablation profiles, the Munnerlyn equation can be shown (Appendix 2.A) to be

$$l(r) = \sqrt{R_1^2 - r^2} - R_1 + \frac{(n-1)R_1}{n-1+SR_1} - \sqrt{\left[\frac{(n-1)R_1}{n-1+SR_1}\right]^2 - r^2}. \quad (2.2)$$

For hyperopic ablation profiles, the transition zone is critically necessary, otherwise there will be a sharp edge at the periphery of the ablation causing a huge curvature change. Such a huge change in corneal curvature is guaranteed to cause night-vision problems when the pupil enlarges at night. Furthermore, a healing effect smoothes this abrupt transition and partially reverses the treatment. Because the transition zone for hyperopic eyes requires a large change in the depth, it requires a wider range. As such, a 6×8 mm zone configuration[3] is often used for myopic ablations, and a 6×9 mm or even 5×9 mm configuration is used for hyperopic ablations.

Example 2.3: Calculate the tissue ablation depth for a +1 D hyperopic eye using a 6 mm optical zone.
Solution: From Eq. (2.2), set $r = 0.006/2$, $S = 1$ D, and use the mean radius of curvature of the cornea as $R_1 = 7.8$ mm, we obtain $l(0.003) = \sqrt{0.0078^2 - (0.006/2)^2} - 0.0078 + \frac{(1.3765-1)\times 0.0078}{1.3765-1+1\times 0.0078} - \sqrt{\left[\frac{(1.3765-1)\times 0.0078}{1.3765-1+1\times 0.0078}\right]^2 - (0.006/2)^2} = 13.50 \times 10^{-6}$ m $= 13.50$ μm. Therefore, for every diopter of hyperopia, the ablation depth is essentially the same as myopia.

2.2.3 Astigmatic Correction

For astigmatic ablation profiles, the design of the Munnerlyn shape is somewhat more complicated. For myopic astigmatic eyes, the myopic profile can be used at different meridians. Because the optical path length is different at different meridians, the transition zone is effectively used to bring the ablation depth to zero at the periphery of the ablation zone. Similarly, for hyperopic astigmatic eyes, the hyperopic profile can be used at different meridians. The ablation profiles for these two types of astigmatism are shown in Fig. 2.5 (a) and (b). For a pure cylinder, the power in one meridian is constant. The ablation profile for a pure cylinder is shown in Fig. 2.5 (c). The optical zone and ablation zone for myopic astigmatism and pure cylinder is the same as those for myopic profile. The optical zone and ablation zone for hyperopic astigmatism is the same as those for the hyperopic profile.

[3] In refractive surgery, laser manufacturers often use OZ×AZ to represent the sizes of the optical zone and the ablation zone. For example, 6×8 mm indicates an optical zone of 6 mm and an ablation zone of 8 mm.

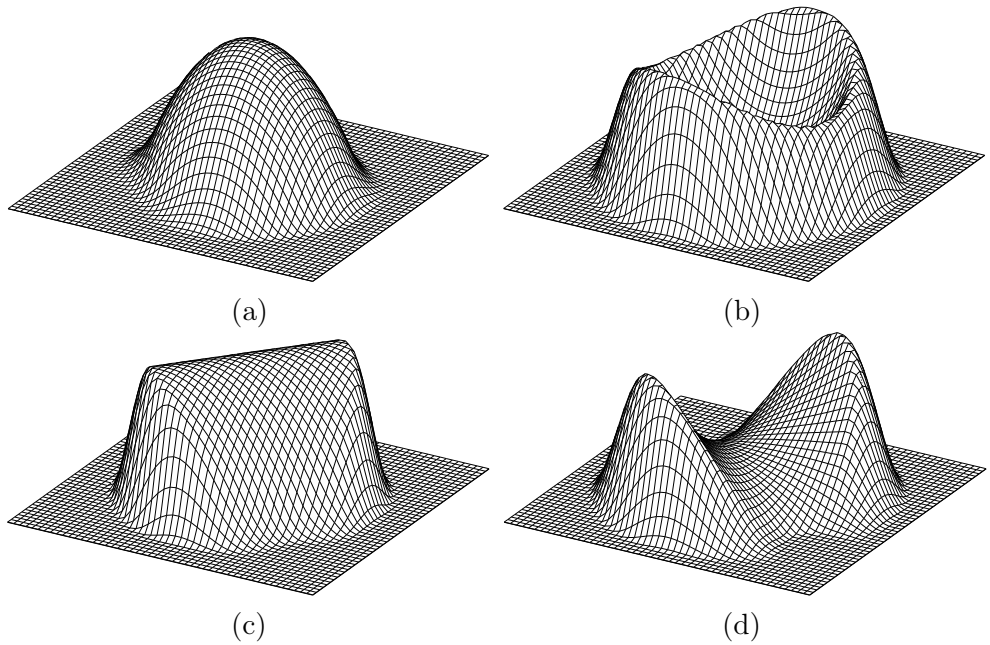


Figure 2.5: Geometry of refractive correction profiles for (a) myopic astigmatism; (b) hyperopic astigmatism; (c) pure cylinder; (d) mixed astigmatism.

A more complex case is the mixed astigmatic ablation profile. In this case, the cylinder power is larger than the sphere power, causing a saddle shape, as shown in Fig. 2.5 (d). The optical path length of the major axis and that of the minor axis are very different. Therefore, the transition zone must be wide enough to account for the smooth transition, so a 6×9 mm configuration is often used.

2.3 Principle of Customized Laser Vision Correction

For laser vision correction, the Munnerlyn equation provides a solution to ablation profiles for spherocylindrical refractive error based on geometrical optics. However, when the high-order ocular aberrations are also aimed for correction, the entire wavefront of the ocular aberrations should be corrected. The basis for the wavefront correction is the phase conjugation, discussed in the previous section.

2.3.1 Wavefront-Driven Correction

When the high-order ocular aberrations are considered, the Munnerlyn equation is not appropriate for calculating the ablation profile. To correct the entire ocular aberrations, including the low-order and the high-order aberrations, the ocular wavefront map may be captured by a wavefront device, such as an aberrometer. With the use of the phase conjugation principle, the ablation profile is simply the inverse of the wavefront map, multiplied by the factor of $1/(n-1)$ to account for the tissue ablation

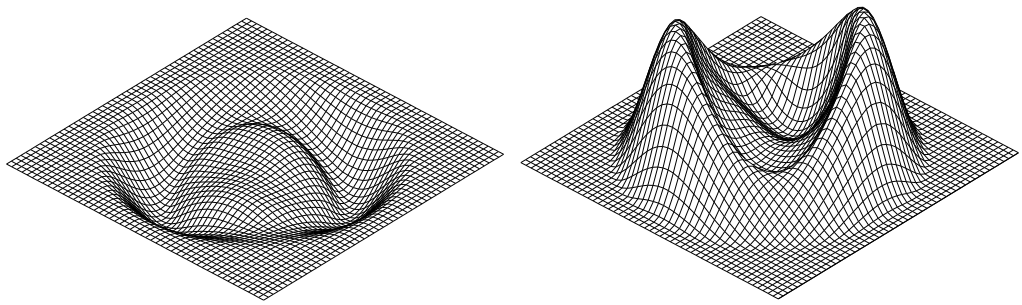


Figure 2.6: Wavefront driven ablation profile. (a) ocular wavefront profile; (b) tissue ablation profile. Note the inversion and magnitude difference between the two profiles.

profile within the optical zone. Similar transition zones as in the Munnerlyn profiles can be used to smooth the transition to zero power along the periphery of the ablation zone. Figure 2.6 (a) shows an example of the wavefront, and Fig. 2.6 (b) shows the corresponding wavefront-driven ablation profile. The inverse of the wavefront shape to the ablation shape can be easily achieved by a sign change. For example, the wavefront indicates a phase lag at one point. This can be seen as a valley in the wavefront map. Therefore, we need to ablate more tissue at that point to increase the optical path length to make up the phase lag. As such, there would be a peak in the ablation profile, so the ablation shape is reversed. The factor of $1/(n-1)$ simply converts the optical path difference to tissue depth.

The mathematical description of the ocular aberrations in wavefront-driven correction can be done by a set of mathematical functions, or basis function, to simplify the analysis. Different sets of functions, such as Zernike polynomials, power series, Taylor monomials, and Fourier series, can be used to represent ocular aberrations. In the next chapter, we will discuss these sets and a special property of Zernike polynomials and Fourier series called orthonormality.

2.3.2 Ablation Equation

When an ocular wavefront is measured with a wavefront device (to be discussed in Chapter 4), the ablation profile can be obtained simply by multiplying the conversion factor $1/(n-1)$ from the wavefront profile to the corneal tissue profile. This tissue profile must be fitted with laser pulses so the desired optical effect of phase conjugation can be achieved after ablation. This problem of arranging laser pulses to produce the desired ablation profile is called the ablation equation, given by

$$W(x,y) = \sum_{i=1}^{N} F_i(x - x_i, y - y_i), \tag{2.3}$$

where N is the total number of laser pulses, F_i is the ablated tissue profile of the

ith laser pulse, and x_i and y_i are the scanning offsets of the ith laser pulse in the x and y directions, respectively. What Eq. (2.3) indicates is that it is possible to use different sizes of laser pulses at difference positions to create an ablation pattern. This is analogous to building a wall with irregular stones. You are allowed to use different sizes of stones to construct a wall of a certain dimension; the fewer gaps between the stones, the better.

If the ablated tissue profile F_i is independent of the sequence number i, the type of the laser is called a fixed-size scanning[4] (FSS) laser. On the other hand, if the profile F_i can be changed from pulse to pulse, the type of the laser is called a variable spot scanning (VSS) laser. The advantage of the VSS laser is the large degree of freedom in selecting the laser pulse sizes to fit an ablation shape. However, the solution of Eq. (2.3) is in general much more difficult for VSS lasers. A more detailed discussion of these two types of lasers is given in Chapter 3.

2.3.3 Profile Fitting with Simulated Annealing

When the ablated tissue profile has fixed shape, Eq. (2.3) can be simplified as

$$W(x, y) = \sum_{i=1}^{N} F(x - x_i, y - y_i). \tag{2.4}$$

Equation (2.4) says that you can only use a laser pulse of one size to make up the ablation profile. With the wall-building analogy, this time only one size stone is used, and therefore it is harder to construct the wall, and the holes between stones are most likely larger. Solution of Eq (2.4) can be achieved by a least square fit. However, for VSS a normal least square fit cannot solve Eq. (2.3) because too many degrees of freedom exist in the equation. Let us make a simple estimation. Suppose we restrict the scanning resolution of the laser pulses to 0.1 mm. For an 8 mm pupil, there can be $80 \times 80 \times \pi/4 \approx 5,000$ possible scanning positions. Let us further assume that the spot size resolution[5] is 0.01 mm, and the spot size ranges between 0.5 mm and 6.5 mm, resulting in 600 possible spot sizes. Therefore, the total degrees of freedom is as high as $5,000 \times 600 = 3,000,000$. With so many degrees of freedom, a sophisticated optimization algorithm must be used.

The method of simulated annealing[7] has long been used to solve large-scale, multidimensional optimization problems. This algorithm is particularly effective in searching for a global minimum when there are a number of local minima in the search space. The technique has been used in the semiconductor industry for integrated circuit design for several decades.

Simulated annealing was inspired and named by the annealing technique in metallurgy, where metals are heated to a high temperature and subsequently cooled over

[4]Most of the fixed-size scanning lasers sometimes are called flying small spot lasers, with the same acronym FSS, or small spot scanning, with the acronym SSS. There is another type of fixed-size scanning laser that uses a rotating slit instead of flying spots.

[5]The actual resolution of the VSS lasers can be much higher than 0.01 mm, because it is practically limited by the moving resolution of the motor to control the size of the iris.

Table 2.1: Pseudo code of the simulated annealing algorithm.

```
s := s0; e := E(s); t := t0     // Init state, energy, temperature
i := 0                          // Count
while i < imax and e > emax     // Downhill in general
  sn := neighbor(s)             // Pick some neighbor
  en := E(sn)                   // New energy
  tn := factor * t              // New temperature
  delta := E(s) - E(sn)         // Energy reduction
  P := exp(delta/tn)            // Current probability
  if random() < P then          // Possible uphill accepted
    s := sn; e := en; t := tn   // Yes, change state
    i := i + 1                  // Increment
  end
end
return s                        // Return value
```

a controlled time period. The heat causes the atoms to deviate from their initial positions and to wander randomly to a new state of higher energy. When the metal is cooled slowly, the atoms can take time to line up orderly to the state of minimum energy. This process can reduce the defects in the metal. On the other hand, if the metal is quenched, or cooled rapidly, the atoms will not be able to line up orderly to reach the state of minimum energy, weakening the structure of the metal.

Table 2.1 shows a pseudo code for the simulated annealing algorithm. Starting from an initial state, the initial energy and temperature are obtained. Then a random selection from the "nearby" solution gives the new energy. The new temperature is obtained by choosing a temperature reduction factor. In general, the algorithm runs in a downhill direction, i.e., the lower energy state. However, it also allows for uphill movements to prevent from getting stuck at the local minima, as long as the probability is lower than the Boltzmann probability distribution. This technique for allowing occasional uphill steps is known as the Metropolis algorithm.[8]

The adoption of the simulated annealing algorithm to solving the ablation equation by the laser manufacturer of VSS type lasers has proven fruitful. With this built-in algorithm to VSS, this type of laser is capable of creating very complex ablation shapes that are essential for wavefront-driven customized refractive surgery.[9]

2.4 Principle of Excimer Laser Ablation of the Cornea

Once an ablation profile is precisely represented by means of a series of laser pulses, as given by Eq. (2.3), the next step is for the laser to deliver these pulses. Excimer lasers are used to deliver such pulses to ablate the corneal tissue. In general, the